

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2002-91473

(P2002-91473A)

(43) 公開日 平成14年3月27日 (2002.3.27)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	テマコード <sup>*</sup> (参考)
G 1 0 L 13/00		G 0 6 F 3/16	3 4 0 A 5 B 0 7 5
G 0 6 F 3/16	3 4 0		3 4 0 N 5 D 0 1 5
		13/00	5 4 0 A 5 D 0 4 5
13/00	5 4 0	17/30	1 1 0 F 5 K 1 0 1
17/30	1 1 0		1 7 0 Z

審査請求 未請求 請求項の数11 O L (全 11 頁) 最終頁に続く

(21) 出願番号 特願2001-178781(P2001-178781)

(22) 出願日 平成13年6月13日 (2001.6.13)

(31) 優先権主張番号 特願2000-197931(P2000-197931)

(32) 優先日 平成12年6月30日 (2000.6.30)

(33) 優先権主張国 日本 (J P)

(71) 出願人 000005223

富士通株式会社

神奈川県川崎市中原区上小田中4丁目1番1号

(72) 発明者 鉄本 秀夫

愛知県名古屋市中区錦1丁目10番1号 株式会社富士通中部システムズ内

(74) 代理人 100092152

弁理士 服部 毅彦

Fターム(参考) 5B075 KK07 ND03 PP10 PP13 PP30

PQ04 PQ36 QM08 UJ40

5D015 KK02 LL05

5D045 AA20 AB26

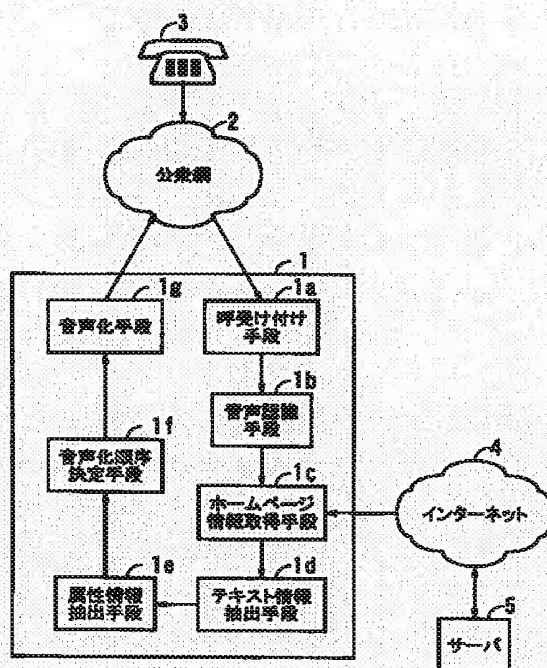
5K101 KK02 LL01 NN08 NN16

(54) 【発明の名称】 情報処理装置

(57) 【要約】

【課題】 ホームページ情報を音声化する際に、文字の大きさ等によって表現されている情報の欠落を防止する。

【解決手段】 呼受け付け手段1aは、ユーザからの電話機3による呼を受け付ける。音声認識手段1bは、電話機3から送信されてきたユーザの音声进行認識する。ホームページ情報取得手段1cは、音声認識手段1bによって所定のホームページにアクセスする要求がなされたことが認識された場合には、対応するホームページ情報を取得する。テキスト情報抽出手段1dは、ホームページ情報に含まれているテキスト情報を、ブロック単位で抽出する。属性情報抽出手段1eは、テキスト情報抽出手段1dによって抽出された各ブロックの属性情報を抽出する。音声化順序決定手段1fは、属性情報抽出手段1eによって抽出された各ブロックの属性情報に応じて、各ブロックの音声化の順序を決定する。音声化手段1gは、音声化順序決定手段1fの決定結果に基づいて、各ブロックに含まれるテキストを音声化する。



## 【特許請求の範囲】

【請求項1】 マークアップ言語で記述されたホームページ情報を読み込んでユーザに提供する処理を行う情報処理装置において、

ユーザからの電話機による呼を受け付ける呼受け付け手段と、

前記電話機から送信されてきたユーザの音声を認識する音声認識手段と、

前記音声認識手段によって所定のホームページにアクセスする要求がなされたことが認識された場合には、対応するホームページ情報を取得するホームページ情報取得手段と、

前記ホームページ情報に含まれているテキスト情報を、ブロック単位で抽出するテキスト情報抽出手段と、

前記テキスト情報抽出手段によって抽出された各ブロックの属性情報を抽出する属性情報抽出手段と、

前記属性情報抽出手段によって抽出された各ブロックの属性情報に応じて、各ブロックの音声化の順序を決定する音声化順序決定手段と、

前記音声化順序決定手段の決定結果に基づいて、各ブロックに含まれるテキストを音声化する音声化手段と、  
を有することを特徴とする情報処理装置。

【請求項2】 前記属性情報は、文字のフォントサイズを示す情報であることを特徴とする請求項1記載の情報処理装置。

【請求項3】 前記属性情報は、各ブロックに含まれるテキストの文字数であることを特徴とする請求項1記載の情報処理装置。

【請求項4】 マークアップ言語で記述されたホームページ情報を読み込んでユーザに提供する処理をコンピュータに機能させるプログラムを記録したコンピュータ読み取り可能な記録媒体において、  
コンピュータを、

ユーザからの電話機による呼を受け付ける呼受け付け手段、

前記電話機から送信されてきたユーザの音声を認識する音声認識手段、

前記音声認識手段によって所定のホームページにアクセスする要求がなされたことが認識された場合には、対応するホームページ情報を取得するホームページ情報取得手段、

前記ホームページ情報に含まれているテキスト情報を、ブロック単位で抽出するテキスト情報抽出手段、

前記テキスト情報抽出手段によって抽出された各ブロックの属性情報を抽出する属性情報抽出手段、

前記属性情報抽出手段によって抽出された各ブロックの属性情報に応じて、各ブロックの音声化の順序を決定する音声化順序決定手段、

前記音声化順序決定手段の決定結果に基づいて、各ブロックに含まれるテキストを音声化する音声化手段、

として機能させるプログラムを記録したコンピュータ読み取り可能な記録媒体。

【請求項5】 マークアップ言語で記述されたホームページ情報を読み込んでユーザに提供する処理をコンピュータに機能させるプログラムにおいて、  
コンピュータを、

ユーザからの電話機による呼を受け付ける呼受け付け手段、

前記電話機から送信されてきたユーザの音声を認識する音声認識手段、

前記音声認識手段によって所定のホームページにアクセスする要求がなされたことが認識された場合には、対応するホームページ情報を取得するホームページ情報取得手段、

前記ホームページ情報に含まれているテキスト情報を、ブロック単位で抽出するテキスト情報抽出手段、

前記テキスト情報抽出手段によって抽出された各ブロックの属性情報を抽出する属性情報抽出手段、

前記属性情報抽出手段によって抽出された各ブロックの属性情報に応じて、各ブロックの音声化の順序を決定する音声化順序決定手段、

前記音声化順序決定手段の決定結果に基づいて、各ブロックに含まれるテキストを音声化する音声化手段、

として機能させることを特徴とするプログラム。

【請求項6】 マークアップ言語で記述されたホームページ情報を読み込んでユーザに提供する処理を行う情報処理装置において、

ユーザからの電話機による呼を受け付ける呼受け付け手段と、

前記電話機から送信されてきたユーザの音声を認識する音声認識手段と、

前記音声認識手段によって所定のホームページにアクセスする要求がなされたことが認識された場合には、対応するホームページ情報を取得するホームページ情報取得手段と、

前記ホームページ情報に含まれている所定の情報を、他の情報に置換する置換手段と、

前記置換手段によって情報の置換がなされたホームページ情報を音声化する音声化手段と、

を有することを特徴とする情報処理装置。

【請求項7】 前記置換手段は、テキスト情報以外の所定の情報を、テキスト情報に置換し、

前記音声化手段は、前記置換手段によって置換されたテキスト情報を音声化する、ことを特徴とする請求項6記載の情報処理装置。

【請求項8】 前記置換手段は、音声情報以外の所定の情報を、音声情報に置換し、

前記音声情報を再生する再生手段を更に有する、ことを特徴とする請求項6記載の情報処理装置。

【請求項9】 前記置換手段は、前記ホームページ情報



と1対1に対応する管理情報を読み出し、その管理情報に応じて置換処理を実行することを特徴とする請求項6記載の情報処理装置。

【請求項10】 マークアップ言語で記述されたホームページ情報を読み込んでユーザに提供する処理をコンピュータに機能させるプログラムを記録したコンピュータ読み取り可能な記録媒体において、

コンピュータを、

ユーザからの電話機による呼を受け付ける呼受け付け手段、

前記電話機から送信されてきたユーザの音声を認識する音声認識手段、

前記音声認識手段によって所定のホームページにアクセスする要求がなされたことが認識された場合には、対応するホームページ情報を取得するホームページ情報取得手段、

前記ホームページ情報に含まれている所定の情報を、他の情報に置換する置換手段、

前記置換手段によって情報の置換がなされたホームページ情報を音声化する音声化手段、

として機能させるプログラムを記録したコンピュータ読み取り可能な記録媒体。

【請求項11】 マークアップ言語で記述されたホームページ情報を読み込んでユーザに提供する処理をコンピュータに機能させるプログラムにおいて、

コンピュータを、

ユーザからの電話機による呼を受け付ける呼受け付け手段、

前記電話機から送信されてきたユーザの音声を認識する音声認識手段、

前記音声認識手段によって所定のホームページにアクセスする要求がなされたことが認識された場合には、対応するホームページ情報を取得するホームページ情報取得手段、

前記ホームページ情報に含まれている所定の情報を、他の情報に置換する置換手段、

前記置換手段によって情報の置換がなされたホームページ情報を音声化する音声化手段、

として機能させることを特徴とするプログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は情報処理装置に関し、特に、マークアップ言語で記述されたホームページ情報を読み込んでユーザに提供する処理を行う情報処理装置に関する。

【0002】

【従来の技術】近年、インターネットの拡張とコンテンツの充実に伴い、種々の情報をネットワークからダウンロードして利用することが可能となりつつある。

【0003】

【発明が解決しようとする課題】ところで、インターネットを利用するには、例えば、パーソナルコンピュータを購入し、インターネット接続サービスを提供するいわゆるプロバイダと契約する必要があるため、ある程度の出費が必要である。また、パーソナルコンピュータの操作にある程度習熟している必要があるため、全ての人が簡易に利用できる状況にあるとは言い難い。

【0004】特に、インターネットでは視覚的な情報が主であるため、視覚に障害を有する者や、弱視者にとっては必ずしも実用性が高い情報源であるとは言い難い側面がある。

【0005】そこで、このような問題を解決するために、例えば、ホームページに記載されたテキスト情報を音声合成により音声化し、電話回線を通じてユーザの電話機から出力する方法も提案されている（特開平10-184249号公報参照）。

【0006】しかしながら、ホームページに掲載されている情報は、例えば、フォントの大きさなどによって、その重要度等が示されているが、テキストを単に音声化しただけではこれらの情報が喪失してしまうという問題点があった。

【0007】また、視覚のみに依存する情報（例えば、画像等）が表示されている場合、その画像に関する情報は完全に喪失してしまうという問題点があった。本発明はこのような点に鑑みてなされたものであり、インターネット上に存在する情報を、可能な限り情報量を減少させることなく音声化することを可能とする情報処理装置を提供することを目的とする。

【0008】

【課題を解決するための手段】本発明では上記課題を解決するために、図1に示す、マークアップ言語で記述されたホームページ情報を読み込んでユーザに提供する処理を行う情報処理装置1において、ユーザからの電話機3による呼を受け付ける呼受け付け手段1aと、前記電話機3から送信されてきたユーザの音声を認識する音声認識手段1bと、前記音声認識手段1bによって所定のホームページにアクセスする要求がなされたことが認識された場合には、対応するホームページ情報を取得するホームページ情報取得手段1cと、前記ホームページ情報に含まれているテキスト情報を、ブロック単位で抽出するテキスト情報抽出手段1dと、前記テキスト情報抽出手段1dによって抽出された各ブロックの属性情報を抽出する属性情報抽出手段1eと、前記属性情報抽出手段1eによって抽出された各ブロックの属性情報に応じて、各ブロックの音声化の順序を決定する音声化順序決定手段1fと、前記音声化順序決定手段1fの決定結果に基づいて、各ブロックに含まれるテキストを音声化する音声化手段1gと、を有することを特徴とする情報処理装置が提供される。

【0009】ここで、呼受け付け手段1aは、ユーザか

らの電話機3による呼を受け付ける。音声認識手段1bは、電話機3から送信されてきたユーザの音声認識する。ホームページ情報取得手段1cは、音声認識手段1bによって所定のホームページにアクセスする要求がなされたことが認識された場合には、対応するホームページ情報を取得する。テキスト情報抽出手段1dは、ホームページ情報に含まれているテキスト情報を、ブロック単位で抽出する。属性情報抽出手段1eは、テキスト情報抽出手段1dによって抽出された各ブロックの属性情報を抽出する。音声化順序決定手段1fは、属性情報抽出手段1eによって抽出された各ブロックの属性情報に応じて、各ブロックの音声化の順序を決定する。音声化手段1gは、音声化順序決定手段1fの決定結果に基づいて、各ブロックに含まれるテキストを音声化する。

【0010】また、マークアップ言語で記述されたホームページ情報を読み込んでユーザに提供する処理を行う情報処理装置において、ユーザからの電話機による呼を受け付ける呼受け付け手段と、前記電話機から送信されてきたユーザの音声認識する音声認識手段と、前記音声認識手段によって所定のホームページにアクセスする要求がなされたことが認識された場合には、対応するホームページ情報を取得するホームページ情報取得手段と、前記ホームページ情報に含まれている所定の情報を、他の情報に置換する置換手段と、前記置換手段によって情報の置換がなされたホームページ情報を音声化する音声化手段と、を有することを特徴とする情報処理装置が提供される。

【0011】ここで、呼受け付け手段は、ユーザからの電話機による呼を受け付ける。音声認識手段は、電話機から送信されてきたユーザの音声認識する。ホームページ情報取得手段は、音声認識手段によって所定のホームページにアクセスする要求がなされたことが認識された場合には、対応するホームページ情報を取得する。置換手段は、ホームページ情報に含まれている所定の情報を、他の情報に置換する。音声化手段は、置換手段によって情報の置換がなされたホームページ情報を音声化する。

【0012】

【発明の実施の形態】以下、本発明の実施の形態を図面を参照して説明する。図1は、本発明に係る情報処理装置の動作原理を説明する原理図である。この図に示すように、本発明に係る情報処理装置1は、呼受け付け手段1a、音声認識手段1b、ホームページ情報取得手段1c、テキスト情報抽出手段1d、属性情報抽出手段1e、音声化順序決定手段1f、および、音声化手段1gによって構成されている。

【0013】呼受け付け手段1aは、ユーザからの電話機3による「呼」を受け付ける。音声認識手段1bは、電話機3から送信されてきた音声認識する。ホームページ情報取得手段1cは、音声認識手段1bによって所

定のホームページにアクセスする要求がなされたことが認識された場合には、対応するホームページ情報を取得する。

【0014】テキスト情報抽出手段1dは、ホームページ情報に含まれているテキスト情報を、ブロック単位で抽出する。属性情報抽出手段1eは、テキスト情報抽出手段1dによって抽出された各ブロックの属性情報を抽出する。

【0015】音声化順序決定手段1fは、属性情報抽出手段1eによって抽出された各ブロックの属性情報に応じて、それぞれのブロックの音声化の順序を決定する。音声化手段1gは、音声化順序決定手段1fの決定結果に基づいて、各ブロックに含まれるテキストを音声化し、公衆網2を介して電話機3に送信する。

【0016】なお、情報処理装置1は、公衆網2を介して電話機3に接続されている。公衆網2は、電話機3と情報処理装置1との間で音声信号を授受する。電話機3は、ユーザの発話に対応する電気信号に変換し、公衆網2を介して情報処理装置1に送信する。

【0017】インターネット4は、情報処理装置1とサーバ5との間でテキスト、画像、音声等の情報を伝送する。サーバ5は、WEBサーバであり、例えば、HTML等のマークアップ言語で記述されたホームページ情報を、情報処理装置1からの要求に応じて送信する。

【0018】次に、以上の原理図の動作について説明する。いま、ユーザが電話機3のハンドセットをオフフックし、情報処理装置1に付与された電話番号に対して発呼したとすると、公衆網2を介して発呼信号が情報処理装置1に対して送り届けられる。

【0019】呼受け付け手段1aは、電話機3からの呼を受け付ける。その結果、電話機3と情報処理装置1との間で通信回線が閉結され、これらの間で通信が可能になる。

【0020】このようにして通信回線が閉結された後、電話機3側のユーザが、例えば、「〇×社のホームページへ接続」のように発話すると、この音声信号は公衆網2を介して音声認識手段1bに送り届けられる。

【0021】音声認識手段1bは、音声認識処理によりこの発話内容が、〇×社のホームページへの接続要求であることを認知し、ホームページ情報取得手段1cにその旨を通知する。

【0022】ホームページ情報取得手段1cは、インターネット4を介して〇×社のホームページ情報（例えば、HTMLによって記述された情報）を、サーバ5から取得する。

【0023】テキスト情報抽出手段1dは、ホームページ情報取得手段1cが取得したホームページ情報に含まれているテキスト情報を、ブロック単位で抽出する。例えば、取得した〇×社のフロントページ（アクセス時に最初に表示されるページ）が3つのブロック（例えば、



タイトル、メニュー、概要説明)から構成されている場合には、それぞれのブロックが個別に抽出される。

【0024】属性情報抽出手段1eは、テキスト情報抽出手段1dが抽出した各ブロックの属性情報を抽出する。ここで、属性情報とは、そのブロックの文字のフォントのサイズ、文字数、または、含まれているハイパーリンクの個数等であり、属性情報抽出手段1eは、各ブロックからこれらの属性情報を抽出する。

【0025】音声化順序決定手段1fは、属性情報抽出手段1eによって抽出された各ブロックの属性情報を参照し、それぞれのブロックの音声化の順序を決定する。例えば、音声化順序決定手段1fは、フォントのサイズが大きい順に各ブロックの音声化順序を決定する。即ち、サイズが大きいフォントで表示されている内容は、重要度が高いと推定されるので、フォントサイズが大きい順に音声化を行う。

【0026】音声化手段1gは、音声化順序決定手段1fによって決定された音声化順序に従って、ブロックを音声化する。その結果、いまの例では、フォントサイズが大きい順(例えば、タイトル、概要説明、メニューの順)に各ブロックが音声化されることになる。

【0027】音声化されたテキスト情報は、公衆網2を介して電話機3に伝送されるので、ユーザは、アクセス要求を行ったホームページに含まれているテキスト情報を、その重要度に応じた順序で聞くことができる。

【0028】以上に説明したように、本発明に係る情報処理装置によれば、インターネット上に存在するホームページ情報を、ユーザからの音声による要求に応じて取得し、取得したホームページに含まれているテキスト情報をブロック単位で抽出し、それぞれのブロックの属性情報に応じた順序で音声化するようにしたので、ユーザは、ホームページに含まれている情報をその重要度に応じた順序で聞くことが可能となる。

【0029】なお、以上の原理図では、フォントのサイズに応じてブロックの音声化順序を決定するようにしたが、例えば、ブロックに含まれているテキストの文字数に応じて音声化順序を決定することも可能である。

【0030】次に、本発明の実施の形態について説明する。図2は、本発明の実施の形態の構成例を示す図である。この図において、電話機10は、ユーザ側に設置され、ユーザの発話内容を対応する電気信号に変換し、公衆網11を介して情報処理装置12に送信するとともに、情報処理装置12から送信されてきた音声信号を、対応する音声に変換して出力する。

【0031】情報処理装置12は、公衆網11を介して電話機10から発呼がなされた場合には電話機10との間で通信回線を閉結し、ユーザからの音声による要求に応じたホームページ情報をサーバ17から取得し、所定の処理を施した後、音声化し、電話機10に対して送信する。

【0032】インターネット16は、サーバ17と情報処理装置12の間で、例えば、HTTP(Hyper Text Transfer Protocol)により、テキスト、画像、音声等からなるホームページ情報を送受信する。

【0033】サーバ17は、WEBサーバであり、HTML等で記述されたホームページ情報を記録しており、情報処理装置12からの要求に応じて、該当するページを読み出して提供する。

【0034】図3は、図2に示す情報処理装置12の詳細な構成例を示す図である。この図に示すように、情報処理装置12は、大別して、電話機10との間の処理を行う音声応答部13、ホームページ情報をダウンロードする制御を行うブラウジング部14、および、ダウンロードしたホームページ情報を解析するHTML解析部15によって構成されている。

【0035】ここで、音声応答部13は、音声認識部13a、ダイヤル認識部13b、および、音声合成部13cによって構成されている。音声認識部13aは、電話機10からの音声信号を認識し、電話操作解析部14aに認識結果を通知する。

【0036】ダイヤル認識部13bは、電話機10のダイヤルが操作された場合には、その操作内容を認識し、電話操作解析部14aに通知する。音声合成部13cは、音声再生制御部14bの制御に従い、出力部15dから供給されるテキスト情報を該当する音声信号に変換し、公衆網11を介して電話機10に対して送信する。

【0037】また、ブラウジング部14は、電話操作解析部14a、音声再生制御部14b、ハイパーリンク制御部14c、および、同一URL内制御部14dによって構成されている。

【0038】電話操作解析部14aは、ユーザが音声またはダイヤル操作によって行った要求を解析し、その解析結果を音声再生制御部14b、ハイパーリンク制御部14c、および、同一URL内制御部14dに通知する。

【0039】音声再生制御部14bは、音声合成部13cによる音声合成処理を制御する。ハイパーリンク制御部14cは、サーバ17に対して所定のホームページ情報の送信を要求する。

【0040】同一URL内制御部14dは、同一のホームページ内において、所定の行から次の行への移動や、他の段落への移動等を制御する。また、HTML解析部15は、構成要素解析部15a、バリアフリーHTML解析部15b、置換部15c、および、出力部15dによって構成されている。

【0041】構成要素解析部15aは、ホームページ情報の構成要素を解析する。バリアフリーHTML解析部15bは、ホームページ情報に拡張されたHTMLとしてのタグが含まれている場合にはこれを解析し、解析結果を構成要素解析部15aに通知する。

【0042】置換部15cは、ホームページ情報に含まれている所定の情報を、他の情報によって置換する処理を実行する。出力部15dは、構成要素解析部15aおよび置換部15cから供給されたホームページ情報を、音声合成部13cに対して供給する。

【0043】次に、本発明の実施の形態の動作について説明する。図4は、情報処理装置12が電話機10からの呼を受けて回線を閉結し、所定の処理を実行した後、回線を閉結するまでの処理の流れを説明するフローチャートである。このフローチャートが開始されると、以下の処理が実行される。

【S1】情報処理装置12は、ユーザからの呼を受けた（着呼した）場合には、ステップS2に進み、それ以外の場合には同一の処理を繰り返す。

【S2】電話操作解析部14aは、例えば、ユーザがダイヤルを操作することにより送信されたパスワードを受信し、正当なユーザであるか否かのユーザ認証を行う。

【0044】なお、この認証処理は、省略することも可能である。

【S3】音声認識部13aがユーザの音声を入力した場合にはステップS4に進み、それ以外の場合には同一の処理を繰り返す。

【S4】音声認識部13aは、受信した音声の認識処理を実行する。

【S5】ブラウジング部14は、ステップS4における認識結果に該当する処理を実行する。なお、具体的には、所定のホームページへの接続要求がなされた場合には、ハイパーリンク制御部14cが該当するホームページのダウンロードを実行する。

【S6】情報処理装置12は処理を終了するか否かを判定し、終了する場合にはステップS7に進み、それ以外の場合にはステップS3に戻って同様の処理を繰り返す。

【0045】例えば、ユーザが電話機10のハンドセットをオンフックして通信回線を切断したか否かを判定し、切断した場合には処理を終了するとして、ステップS7に進む。

【S7】情報処理装置12は、電話機10との間に閉結された通信回線を切断する処理を実行する。

【0046】以上の処理によれば、ユーザが電話機10を介して発話した内容や、ダイヤルの操作内容に応じた処理を、情報処理装置12に実行させることが可能となる。次に、図5を参照して本発明の第1の実施の形態の動作について説明する。本発明の第1の実施の形態では、ホームページ情報に含まれているテキスト情報を、ブロック単位で抽出し、各ブロックの属性情報に応じた順序で、それぞれのブロックを音声化する。このフローチャートが開始されると、以下の処理が実行される。

【S20】ハイパーリンク制御部14cは、ユーザが音声によって要求した所定のホームページ情報を、サーバ

17に要求する。

【0047】例えば、ハイパーリンク制御部14cは、ユーザからの要求に応じて、図6に示すようなホームページ情報をサーバ17から取得するように要求する。なお、この例では、3つのセルからなる表がウィンドウ50の表示領域50aに表示されており、左端のセルは上下2つのセルに更に分割されている。また、左端の上側のセル50bには「選挙公示、立候補者1400人」が表示されており、その下のセル50cには「選挙が公示された。午後4時・・・」が表示されている。中央のセルには、画像が表示され、また、右端のセル50dには、メニュー項目が列記されている。

【0048】【S21】構成要素解析部15aは、ハイパーリンク制御部14cの要求に応じて送信されてきたホームページ情報を入力し、タグによって圍繞された要素としてのエレメントに展開する。なお、エレメントはHTML文書における意味のかたまりを示し、また、エレメントは階層構造を有しているため、適当な階層によってHTML文書を分割した場合には、各エレメントはブロックと等しくなる。

【0049】図7は、図6に対応するHTML文書の一例を示す図である。この図に示すように、図6に示すホームページに対応するHTML文書は、属性を示すタグとテキスト情報とが組み合わされて形成されている。

【0050】構成要素解析部15aは、このようなホームページ情報を、エレメント毎に展開する。ここで図6の例では、「選挙公示、立候補・・・」と、「選挙が公示された。午後4時・・・」と、「●南北首脳会談・・・」とがそれぞれ異なるセルに格納されているので、別個のエレメントと認識される。

【S22】構成要素解析部15aは、全エレメントに対する処理が完了したか否かを判定し、完了した場合には処理を終了する。また、それ以外の場合には、ステップS23に進む。

【S23】構成要素解析部15aは、HTML文書に含まれているタグを解析する。

【0051】即ち、構成要素解析部15aは、それぞれのエレメントに属するタグに埋め込まれているフォントサイズ等の属性情報を解析する。具体的には、図6に示すタイトル「選挙公示、立候補・・・」の場合では、図7の第6行目に示すように、そのフォントサイズは、6（font size="6"）であるので、これが取得される。

【S24】構成要素解析部15aは、それぞれのエレメントに含まれているテキストを解析する。

【0052】即ち、構成要素解析部15aは、それぞれのエレメントに含まれているテキストを解析し、その文字数等を取得する。

【S25】構成要素解析部15aは、ステップS23およびステップS24の解析結果に応じて、そのエレメン

10

20

30

40

50



トの読み上げ順位を決定し、エレメントテーブルに格納する。ここで、図6の例では、タイトル「選挙公示、立候補者・・・」のフォントサイズが“6”で最大であるので、このタイトルの順位が“1”となる。また、テキスト「選挙が公示された。午後4時・・・」およびメニュー「●南北首脳会談・・・」のフォントサイズは、それぞれ“3”、“2”であるので、それぞれの順位は“2”、“3”となる。

【0053】図8は、図6に示すホームページに対応するエレメントテーブルの一例である。この例では、図6の例を構成する各エレメントが、順位とともに格納されている。例えば、第1番目の項目は、図6の左上のタイトル「選挙公示、立候補者・・・」に対応しており、その順位は“1”とされている。

【0054】以上の処理により、ホームページ情報を構成するエレメントに対して、そのエレメントに含まれるテキスト情報の属性に応じた読み上げ順序が付与される。このようにして生成されたエレメントテーブルは、テキストデータとともに、出力部15dを介して音声合成部13cに供給される。

【0055】音声合成部13cは、エレメントテーブルを参照し、供給されたテキストデータを音声化する処理を実行する。いまの例では、エレメントテーブルにおいて順位が“1”である、図6に示す左端のセル50bのタイトルである「選挙公示、立候補・・・」が最初に音声化される。続いて、その下のセル50cの「選挙が公示された。午後4時現在・・・」が音声化され、最後に右端のセル50dの「●南北首脳会談・・・」が音声化されることになる。

【0056】以上に説明したように、本発明の実施の形態によれば、ホームページを構成するテキスト情報を、ブロック（エレメント）単位で分割し、それぞれのブロックに含まれているテキスト情報の属性に応じて、音声化の順番を決定するようにしたので、重要な情報を優先して音声化することが可能となる。

【0057】なお、以上の実施の形態では、フォントサイズを基準にして、各ブロックの音声化順序を決定するようにしたが、例えば、ブロックに含まれているテキストの文字数その他を基準にするようにしてもよい。

【0058】また、以上の実施の形態では、表のセルに含まれているテキスト情報を1つのブロック（エレメント）として扱うようにしたが、このような方法のみならず、例えば、隣接して配置されているテキスト情報は、1つのブロックとみなすようにしてもよい。

【0059】次に、本発明の第2の実施の形態について説明する。図9は、本発明の第2の実施の形態の動作を説明するフローチャートである。なお、第2の実施の形態の構成は、第1の実施の形態の場合と同様であるのでその説明は省略する。

【0060】このフローチャートが開始されると、以下

の処理が実行される。

【S40】ハイパーリンク制御部14cは、ユーザが音声によって要求した所定のホームページ情報を、サーバ17から取得するように要求する。

【0061】例えば、ハイパーリンク制御部14cは、ユーザからの要求に応じて、図10に示すようなホームページ情報をサーバ17から取得するように要求する。なお、この例では、タイトル「今週のヒットチャート」が表示されるとともに、ヒットチャートにランクインした曲名と歌手の名前が第1位から第4位まで表示されている。

【S41】構成要素解析部15aは、ハイパーリンク制御部14cの要求に応じて送信されてきたホームページ情報を入力し、タグによって囲繞された要素としてのエレメントに展開する。

【0062】即ち、構成要素解析部15aは、図11に示すような対象となるホームページのHTML文書を取得し、このHTML文書に含まれているエレメント（タグによって囲繞された構成要素）を抽出し、エレメント

テーブルを生成する。

【0063】図12は、生成されたエレメントテーブルの一例を示す図である。この例では、各エレメントに含まれているデータが、エレメント番号、サイズ、属性等と共に格納されている。

【S42】置換部15cは、構成要素解析部15aから解析結果を受け取るとともに、図示せぬ記憶部から該当するファイル置換テーブルを読み込む。

【0064】ここで、ファイル置換テーブルとは、ホームページに含まれている情報を他の情報に置換したり、ホームページに含まれている情報に新たな情報を付加するためのテーブルであり、図13にその一例を示す。この例では、図12に示す各エレメントに対応する曲が記録されたファイルのファイル名がエレメント番号に対応付けて格納されている。例えば、エレメント番号「EL E001」に対応する置換データは、「grand\_son. mp3」であり、そのファイル形式は拡張子から「MP3」形式であることが分かる。

【S43】置換部15cは、ファイル置換テーブルを参照し、該当するエレメントのデータを置換する。

【0065】例えば、第1番目のエレメントの場合では、エレメントに含まれているデータ「1位 GRAND SON・・・」が、MP3形式の音楽データである grand\_son. mp3に置換される。なお、置換せずに、MP3形式のファイルを付加することも可能である。

【S44】音声合成部13cは、置換がなされたホームページ情報の供給を受け、テキスト情報を音声化するとともに、置換によって新たに付加された音楽データがある場合には、その音楽データを再生する。

【0066】従って、以上の処理によって置換がなされ

た図10に示すホームページを音声化した場合、先ず、ウィンドウ60の表示領域60aに表示されている「今週のヒットチャート」が発話された後、ヒットチャートの第1位から第4位にランクインしている曲が順に再生されることになる。

【0067】なお、以上の実施の形態では、エレメントに含まれる一部のデータを音楽データに置換するようにしたが、置換せずにエレメントの末尾に音楽データを付加するようにしてもよい。そのような構成によれば、

「1位 GRAND SON浅田明」が発話された後、音楽ファイル「grand\_son.mp3」が再生されることになる。

【0068】また、以上の実施の形態では、ホームページ単位でファイル置換テーブルを準備し、サーバからホームページ情報が取得された場合には、該当するファイル置換テーブルを取得して置換処理を行うようにしたが、置換対象となるデータを、HTMLの拡張タグによって指示するようにしてもよい。

【0069】例えば、HTMLのタグを拡張してバリアフリータグ<BF>を定義し、このタグに対して追加しようとするファイルのファイル名を埋め込んでおくようにしてもよい。なお、このようにして埋め込まれたバリアフリータグは、通常のブラウザで閲覧する場合には、無意味なタグとして無視されるが、ポータルとしての情報処理装置12に対して、例えば、バリアフリーHTML解析部15bを具備しておき、このバリアフリーHTML解析部15bにより、バリアフリータグ<BF>を解析するようにすることもできる。

【0070】更に、以上の実施の形態では、テキスト情報を音声情報に置換するようにしたが、例えば、画像情報を直接音声情報に置換したり、画像情報を該当するテキスト情報（例えば、その画像を説明するテキスト情報）に置換し、音声合成部13cによって音声化するようにしてもよい。

【0071】最後に、上記の処理機能は、コンピュータによって実現することができる。その場合、情報処理装置が有すべき機能の処理内容は、コンピュータで読み取り可能な記録媒体に記録されたプログラムに記述されており、このプログラムをコンピュータで実行することにより、上記処理がコンピュータで実現される。コンピュータで読み取り可能な記録媒体としては、磁気記録装置や半導体メモリ等がある。市場へ流通させる場合には、CD-ROM(Compact Disk Read Only Memory)やフレキシブルディスク等の可搬型記録媒体にプログラムを格納して流通させたり、ネットワークを介して接続されたコンピュータの記憶装置に格納しておき、ネットワークを通じて他のコンピュータに転送することもできる。コンピュータで実行する際には、コンピュータ内のハードディスク装置等にプログラムを格納しておき、メインメモリにロードして実行する。

【0072】

【発明の効果】以上説明したように本発明では、マークアップ言語で記述されたホームページ情報を読み込んでユーザに提供する処理を行う情報処理装置において、ユーザからの電話機による呼を受け付ける呼受け付け手段と、電話機から送信されてきたユーザの音声認識する音声認識手段と、音声認識手段によって所定のホームページにアクセスする要求がなされたことが認識された場合には、対応するホームページ情報を取得するホームページ情報取得手段と、ホームページ情報に含まれているテキスト情報を、ブロック単位で抽出するテキスト情報抽出手段と、テキスト情報抽出手段によって抽出された各ブロックの属性情報を抽出する属性情報抽出手段と、属性情報抽出手段によって抽出された各ブロックの属性情報に応じて、各ブロックの音声化の順序を決定する音声化順序決定手段と、音声化順序決定手段の決定結果に基づいて、各ブロックに含まれるテキストを音声化する音声化手段と、を有するようにしたので、ホームページ情報を音声化する際に、文字の大きさ等の属性によって表現されているホームページ上の情報を、発話順序によって表すことが可能となる。

【0073】また、マークアップ言語で記述されたホームページ情報を読み込んでユーザに提供する処理を行う情報処理装置において、ユーザからの電話機による呼を受け付ける呼受け付け手段と、電話機から送信されてきたユーザの音声認識する音声認識手段と、音声認識手段によって所定のホームページにアクセスする要求がなされたことが認識された場合には、対応するホームページ情報を取得するホームページ情報取得手段と、ホームページ情報に含まれている所定の情報を、他の情報に置換する置換手段と、置換手段によって情報の置換がなされたホームページ情報を音声化する音声化手段と、を有するようにしたので、例えば、画像情報を該当する音声情報に置換することにより、ホームページ情報を音声化する際に、情報の欠落を防止することができる。

【図面の簡単な説明】

【図1】本発明の動作原理を説明する原理図である。

【図2】本発明の実施の形態の構成例を示す図である。

【図3】図2に示す情報処理装置の詳細な構成例を示す図である。

【図4】図3に示す実施の形態において、ユーザからの着呼から切断処理までの処理の流れを説明するためのフローチャートである。

【図5】ブロックの音声化順序を決定するための処理の流れを説明するためのフローチャートである。

【図6】図5に示すフローチャートが処理の対象とするホームページの一例である。

【図7】図6に示すホームページに対応するHTML文書である。

【図8】図5に示す処理を図6に示すホームページに適



用した場合に得られるエレメントテーブルの一例である。

【図9】本発明の第2の実施の形態の動作を説明するためのフローチャートである。

【図10】図9に示すフローチャートが処理の対象とするホームページの一例である。

【図11】図10に示すホームページに対応するHTML文書である。

【図12】エレメントテーブルの一例を示す図である。

【図13】ファイル置換テーブルの一例を示す図である。

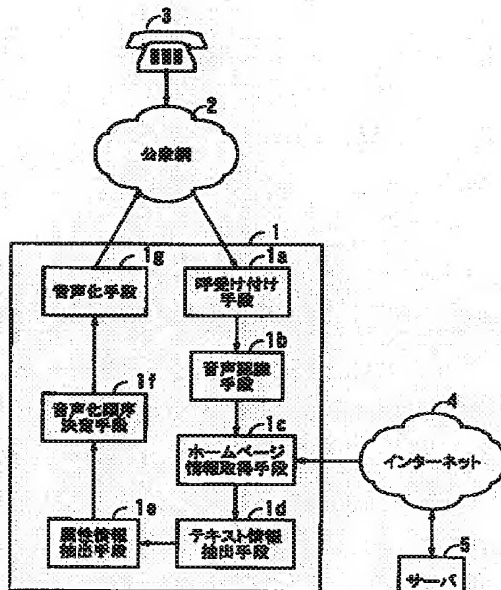
【符号の説明】

- 1 情報処理装置
- 1a 呼受け付け手段
- 1b 音声認識手段
- 1c ホームページ情報取得手段
- 1d テキスト情報抽出手段
- 1e 属性情報抽出手段
- 1f 音声化順序決定手段
- 1g 音声化手段
- 2 公衆網
- 3 電話機

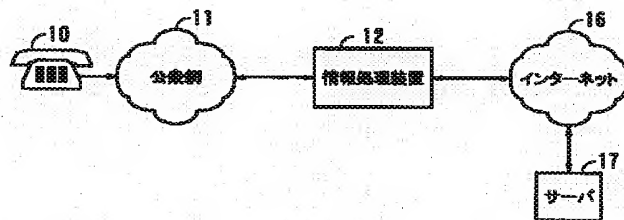
- \* 4 インターネット
- 5 サーバ
- 10 電話機
- 11 公衆網
- 12 情報処理装置
- 13 音声応答部
- 13a 音声認識部
- 13b ダイアル認識部
- 13c 音声合成部
- 14 ブラウジング部
- 14a 電話操作解析部
- 14b 音声再生制御部
- 14c ハイパーリンク制御部
- 14d 同一URL内制御部
- 15 HTML解析部
- 15a 構成要素解析部
- 15b バリアフリーHTML解析部
- 15c 置換部
- 15d 出力部
- 20 16 インターネット
- 17 サーバ

\*

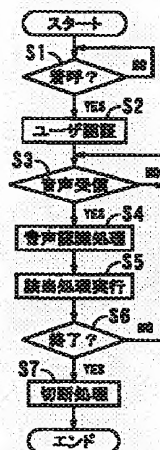
【図1】



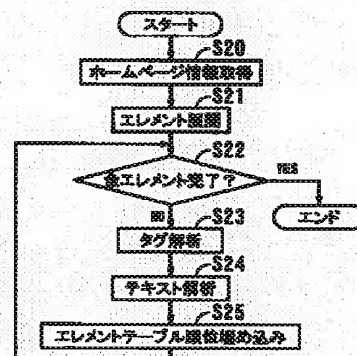
【図2】



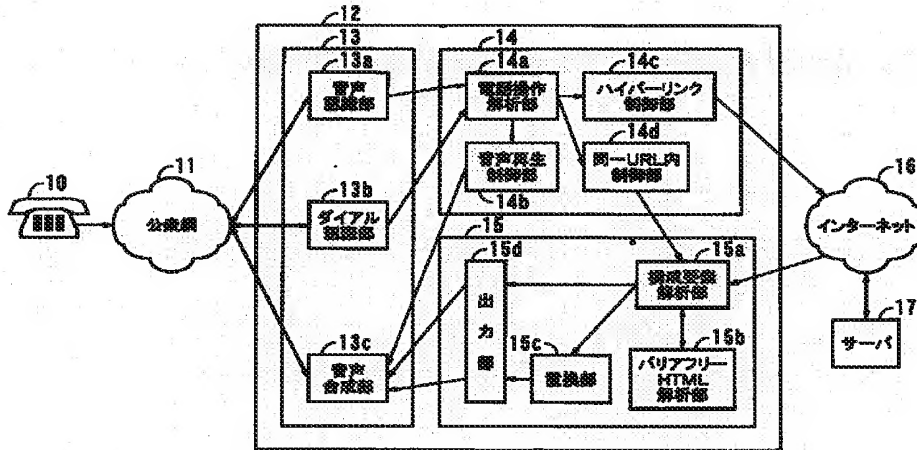
【図4】



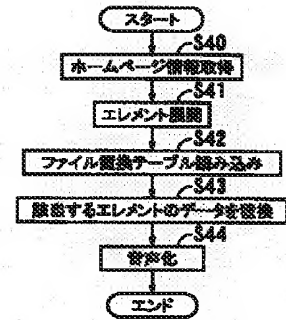
【図5】



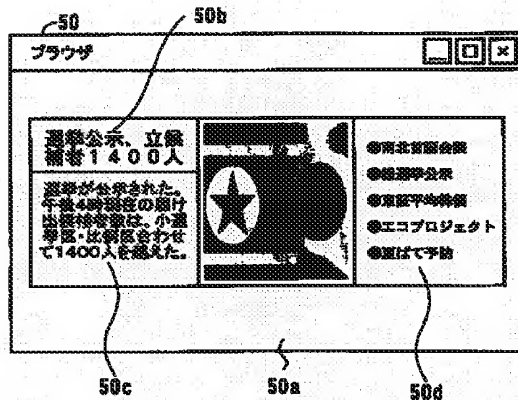
【図3】



【図9】



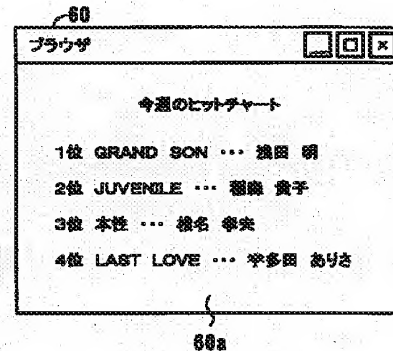
【図6】



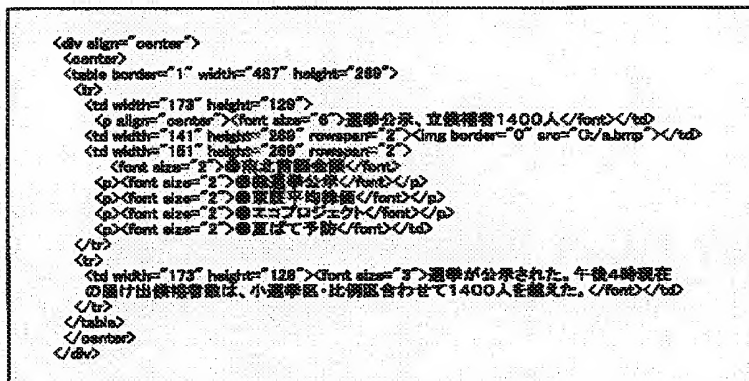
【図8】

エレメント番号	サイズ	属性	データ	順位
ELE001	28バイト	テキスト	選挙公示、立候補者...	1
ELE002	94バイト	テキスト	選挙が公示された...	2
ELE003	72バイト	テキスト	南北両派会派...	3

【図10】



【図7】



【図13】

エレメント番号	置換データ
ELE001	grand_son. mp3
ELE002	juvenile. mp3
ELE003	dharma. mp3
ELE004	last_love. mp3



【図11】

```

<body>
<h1 align="center">今週のヒットチャート</h1>
<p align="center"></p>
<p align="left">1位 GRAND SON ... 渡辺研</p>
<p align="left">2位 JUVENILE ... 藤崎 貴子</p>
<p align="left">3位 本性 ... 椎名 幸夫</p>
<p align="left">4位 LAST LOVE ... 宇多田 ありき</p>
</body>

```

【図12】

エレメント番号	サイズ	属性	データ
ELE001	18バイト	テキスト	1位 GRAND ...
ELE002	18バイト	テキスト	2位 JUVENILE ...
ELE003	4バイト	テキスト	3位 本性 ...
ELE004	18バイト	テキスト	4位 LAST ...

## フロントページの続き

(51)Int.Cl.	識別記号	F I	キーワード (参考)
G 0 6 F 17/30	1 7 0	G 0 6 F 17/30	3 1 0 Z
	3 1 0		3 6 0 Z
	3 6 0	H 0 4 M 11/00	3 0 2
G 1 0 L 15/00		G 1 0 L 3/00	E
15/22			5 5 1 P
15/28			5 6 1 D
21/06			R
H 0 4 M 11/00	3 0 2		S

**\* NOTICES \***

**JPO and INPIT are not responsible for any damages caused by the use of this translation.**

1.This document has been translated by computer. So the translation may not reflect the original precisely.

2.\*\*\* shows the word which can not be translated.

3.In the drawings, any words are not translated.

---

**CLAIMS**

---

[Claim(s)]

[Claim 1]An information processor which performs processing which reads homepage information described by a markup language, and with which a user is provided, comprising:

A call receptionist means to receive a call by telephone from a user.

A voice recognition means which recognizes a sound of a user transmitted from said telephone.

A homepage information acquisition means which acquires corresponding homepage information when it has been recognized that a demand which accesses a predetermined homepage by said voice recognition means was made.

A text information extracting means which extracts contained-text information by a block unit to said homepage information.

An attribute information extracting means which extracts attribution information of each block extracted by said text information extracting means.

A sound-ized order determination means which determines an order of sound-izing of each block according to attribution information of each block extracted by said attribute information extracting means.

A sound-ized means to sound-size a text contained in each block based on decision results of said sound-ized order determination means.

[Claim 2]The information processor according to claim 1, wherein said attribution information is information which shows a font size of a character.

[Claim 3]The information processor according to claim 1, wherein said attribution



information is the number of characters of a text contained in each block.

[Claim 4]In a recording medium for which a program which operates as a computer processing which reads homepage information described by a markup language, and with which a user is provided was recorded and in which computer reading is possible, A call receptionist means to receive a call according a computer to telephone from a user, When it has been recognized that a demand which accesses a predetermined homepage by voice recognition means which recognizes a sound of a user transmitted from said telephone, and said voice recognition means was made, Corresponding homepage information to a homepage information acquisition means to acquire and said homepage information contained-text information, A text information extracting means extracted by a block unit, an attribute information extracting means which extracts attribution information of each block extracted by said text information extracting means, According to attribution information of each block extracted by said attribute information extracting means, A recording medium which recorded a program which makes a text contained in each block a sound-sized means to sound-size, and as which it is operated based on decision results of a sound-sized order determination means which determines an order of sound-izing of each block, and said sound-sized order determination means and in which computer reading is possible.

[Claim 5]In a program which operates as a computer processing which reads homepage information described by a markup language, and with which a user is provided, A call receptionist means to receive a call according a computer to telephone from a user, When it has been recognized that a demand which accesses a predetermined homepage by voice recognition means which recognizes a sound of a user transmitted from said telephone, and said voice recognition means was made, Corresponding homepage information to a homepage information acquisition means to acquire and said homepage information contained-text information, A text information extracting means extracted by a block unit, an attribute information extracting means which extracts attribution information of each block extracted by said text information extracting means, According to attribution information of each block extracted by said attribute information extracting means, A program making a text contained in each block into a sound-sized means to sound-size, and operating it based on decision results of a sound-sized order determination means which determines an order of sound-izing of each block, and said sound-sized order determination means.

[Claim 6]An information processor which performs processing which reads homepage information described by a markup language, and with which a user is provided, comprising:

A call receptionist means to receive a call by telephone from a user.

A voice recognition means which recognizes a sound of a user transmitted from said telephone.

A homepage information acquisition means which acquires corresponding homepage information when it has been recognized that a demand which accesses a predetermined homepage by said voice recognition means was made.

A permutation means which replaces predetermined information included in said homepage information by other information, and a sound-ized means to sound-ize homepage information by which substitution of information was made by said permutation means.

[Claim 7]The information processor according to claim 6 characterized by what said permutation means replaces predetermined information other than text information by text information, and said sound-ized means sound-izes text information replaced by said permutation means for.

[Claim 8]The information processor according to claim 6 characterized by what said permutation means replaces predetermined information other than speech information by speech information, and it has further a reproduction means which reproduces said speech information for.

[Claim 9]The information processor according to claim 6 said permutation means's reading said homepage information and management information corresponding to 1 to 1, and performing substitution processing according to the management information.

[Claim 10]In a recording medium for which a program which operates as a computer processing which reads homepage information described by a markup language, and with which a user is provided was recorded and in which computer reading is possible, A call receptionist means to receive a call according a computer to telephone from a user, When it has been recognized that a demand which accesses a predetermined homepage by voice recognition means which recognizes a sound of a user transmitted from said telephone, and said voice recognition means was made, Predetermined information included in a homepage information acquisition means which acquires corresponding homepage information, and said homepage information, A recording medium which recorded a program which considers it as a sound-ized means to sound-ize homepage information by which substitution of information was made, and is operated by permutation means replaced by other information, and said permutation means and in which computer reading is possible.

[Claim 11]In a program which operates as a computer processing which reads



homepage information described by a markup language, and with which a user is provided, A call receptionist means to receive a call according a computer to telephone from a user, When it has been recognized that a demand which accesses a predetermined homepage by voice recognition means which recognizes a sound of a user transmitted from said telephone, and said voice recognition means was made, Predetermined information included in a homepage information acquisition means which acquires corresponding homepage information, and said homepage information, A program considering it as a sound-ized means to sound-ize homepage information by which substitution of information was made, and making it function by permutation means replaced by other information, and said permutation means.

---

[Translation done.]

---

## DETAILED DESCRIPTION

---

[Detailed Description of the Invention]

[0001]

[Field of the Invention]Especially this invention relates to the information processor which performs processing which reads the homepage information described by the markup language about an information processor, and with which a user is provided.

[0002]

[Description of the Prior Art]It is becoming possible to download and use various information from a network with extension of the Internet, and fullness of contents in recent years.

[0003]

[Problem(s) to be Solved by the Invention]By the way, since it is necessary to contract with what is called a provider that purchases a personal computer and provides an Internet access service, for example in order to use the Internet, a certain amount of expenses are required. Since it needs to be skilled in operation of a personal computer to some extent, it is hard to say that it is in the situation which all people can use simply.

[0004]In particular, on the Internet, since visual information is main, the side which is hard to say that practicality is not necessarily high sources of information for the person and amblyopia person who have an obstacle is in vision.

[0005]Then, in order to solve such a problem, for example, text information indicated to the homepage is sound-ized by voice synthesis, and the method of outputting from a user's telephone through a telephone line is also proposed (refer to JP,10-164249,A).

[0006]However, there was a problem that the information published by the homepage will be lost by these information only by sound-izing a text, for example although the importance is shown by the size of the font, etc.

[0007]When the information only depending on vision, including for example, a picture etc., was displayed, the information about the picture had the problem of being lost thoroughly. This invention is made in view of such a point, and is a thing.

The purpose is to provide the information processor which makes it possible to sound-ize the information which exists upwards, without decreasing the amount of information as much as possible.

[0008]

[Means for Solving the Problem]In the information processor 1 which performs processing which reads homepage information which is shown in drawing 1, and which was described by a markup language, and with which a user is provided in order to solve an aforementioned problem in this invention, A call receptionist means 1a to receive a call by the telephone 3 from a user, The voice recognition means 1b which recognizes a sound of a user transmitted from said telephone 3, When it has been recognized that a demand which accesses a predetermined homepage by said voice recognition means 1b was made, The homepage information acquisition means 1c which acquires corresponding homepage information, The text information extracting means 1d which extracts contained-text information by a block unit to said homepage information, The attribute information extracting means 1e which extracts attribution information of each block extracted by said text information extracting means 1d, The sound-ized order determination means 1f which determines an order of sound-izing of each block according to attribution information of each block extracted by said attribute information extracting means 1e, An information processor having a sound-ized means 1g to sound-ize a text contained in each block, based on decision results of said sound-ized order determination means 1f is provided.

[0009]Here, the call receptionist means 1a receives a call by the telephone 3 from a user. The voice recognition means 1b recognizes a sound of a user transmitted from the telephone 3. The homepage information acquisition means 1c acquires corresponding homepage information, when it has been recognized that a demand



which accesses a predetermined homepage by the voice recognition means 1b was made. The text information extracting means 1d extracts contained-text information by a block unit to homepage information. The attribute information extracting means 1e extracts attribution information of each block extracted by the text information extracting means 1d. The sound-ized order determination means 1f determines an order of sound-izing of each block according to attribution information of each block extracted by the attribute information extracting means 1e. The sound-ized means 1g sound-izes a text contained in each block based on decision results of the sound-ized order determination means 1f.

[0010]In an information processor which performs processing which reads homepage information described by a markup language, and with which a user is provided, A call receptionist means to receive a call by telephone from a user, and a voice recognition means which recognizes a sound of a user transmitted from said telephone, When it has been recognized that a demand which accesses a predetermined homepage by said voice recognition means was made, A homepage information acquisition means which acquires corresponding homepage information, An information processor having a permutation means which replaces predetermined information included in said homepage information by other information, and a sound-ized means to sound-ize homepage information by which substitution of information was made by said permutation means is provided.

[0011]Here, a call receptionist means receives a call by telephone from a user. A voice recognition means recognizes a sound of a user transmitted from telephone. A homepage information acquisition means acquires corresponding homepage information, when it has been recognized that a demand which accesses a predetermined homepage by a voice recognition means was made. A permutation means replaces predetermined information included in homepage information by other information. A sound-ized means sound-izes homepage information by which substitution of information was made by a permutation means.

[0012]

[Embodiment of the Invention]Hereafter, an embodiment of the invention is described with reference to drawings. Drawing 1 is a principle figure explaining the principle of operation of the information processor concerning this invention. As shown in this figure, the information processor 1 concerning this invention is constituted by the call receptionist means 1a, the voice recognition means 1b, the homepage information acquisition means 1c, the text information extracting means 1d, the attribute information extracting means 1e, the sound-ized order determination means 1f, and

the sound-ized means 1g.

[0013]The call receptionist means 1a receives the "call" by the telephone 3 from a user. The voice recognition means 1b recognizes the sound transmitted from the telephone 3. The homepage information acquisition means 1c acquires corresponding homepage information, when it has been recognized that the demand which accesses a predetermined homepage by the voice recognition means 1b was made.

[0014]The text information extracting means 1d extracts contained-text information by a block unit to homepage information. The attribute information extracting means 1e extracts the attribution information of each block extracted by the text information extracting means 1d.

[0015]The sound-ized order determination means 1f determines an order of sound-izing of each block according to the attribute of each block extracted by the attribute information extracting means 1e. The sound-ized means 1g sound-izes the text contained in each block based on the decision results of the sound-ized order determination means 1f, and transmits to the telephone 3 via the public network 2.

[0016]The information processor 1 is connected to the telephone 3 via the public network 2. The public network 2 delivers and receives an audio signal between the telephone 3 and the information processor 1. The telephone 3 changes a user's utterance into a corresponding electrical signal, and transmits to the information processor 1 via the public network 2.

[0017]The Internet 4 transmits information, including a text, a picture, a sound, etc., between the information processor 1 and the server 5. The server 5 is a WEB server, for example, transmits the homepage information described by markup languages, such as HTML, according to the demand from the information processor 1.

[0018]Next, operation of the above principle figure is explained. Now, a user does off-hook [ of the hand set of the telephone 3 ], and supposing it carries out call origination to the telephone number given to the information processor 1, a calling signal will be sent to the information processor 1 via the public network 2.

[0019]The call receptionist means 1a receives the call from the telephone 3. As a result, a communication line is closed between the telephone 3 and the information processor 1, and communication becomes possible among these.

[0020]Thus, if the user of the telephone 3 side speaks like "connecting with the homepage of Ox company", for example after a communication line is closed, this audio signal will be sent into the voice recognition means 1b via the public network 2.

[0021]These contents of utterance recognize that it is a connection request to the homepage of Ox company by speech recognition processing, and the voice recognition

means 1b notifies that to the homepage information acquisition means 1c.

[0022]The homepage information acquisition means 1c acquires the homepage information (for example, information described by HTML) of Ox company from the server 5 via the Internet 4.

[0023]The text information extracting means 1d extracts contained-text information by a block unit to the homepage information which the homepage information acquisition means 1c acquired. For example, when the acquired front page (page first displayed at the time of access) of Ox company comprises three blocks (for example, a title, a menu, outline explanation), each block is extracted individually.

[0024]The attribute information extracting means 1e extracts the attribution information of each block which the text information extracting means 1d extracted. Here, attribution information is the size of the font of the character of the block, the number of characters, or the number of a hyperlink contained, and the attribute information extracting means 1e extracts such attribution information from each block.

[0025]The sound-sized order determination means 1f determines an order of sound-izing of each block with reference to the attribution information of each block extracted by the attribute information extracting means 1e. For example, the sound-sized order determination means 1f determines a sound-sized order of each block of the size of a font in descending. That is, since it is presumed that the contents currently displayed with the font with large size have high importance, a font size performs sound-ization to descending.

[0026]The sound-sized means 1g sound-sizes a block according to the sound-sized order determined by the sound-sized order determination means 1f. As a result, in the present example, each block will be sound-sized for a font size by descending (for example, a title, outline explanation, order of a menu).

[0027]Since the sound-sized text information is transmitted to the telephone 3 via the public network 2, the user can ask contained-text information to the homepage which performed the access request in an order according to the importance.

[0028]As explained above, according to the information processor concerning this invention, the homepage information which exists on the Internet, Since it acquires according to the demand with the sound from a user, contained-text information is extracted by a block unit to the acquired homepage and it was [ sound-] made toize in an order according to the attribution information of each block, A user becomes possible [ hearing the information included in the homepage in an order according to the importance ].

[0029]Although a sound-sized order of the block was determined according to the size



of a font, it is also possible to determine a sound-sized order as a block for example, according to the number of characters of a contained text in the above principle figure.

[0030]Next, an embodiment of the invention is described. Drawing 2 is a figure showing the example of composition of an embodiment of the invention. In this figure, it changes and outputs the audio signal transmitted from the information processor 12 to a corresponding sound while the telephone 10 is installed in the user side, changes a user's contents of utterance into a corresponding electrical signal and transmits to the information processor 12 via the public network 11.

[0031]When call origination is made from the telephone 10 via the public network 11, after the information processor 12 closes a communication line between the telephones 10, acquires the homepage information according to the demand with the sound from a user from the server 17 and performs predetermined processing, it is sound-sized and transmits to the telephone 10.

[0032]The Internet 16 transmits by HTTP (Hyper Text Transfer Protocol) and receives the homepage information which consists of a text, a picture, a sound, etc. between the server 17 and the information processor 12, for example.

[0033]The server 17 is a WEB server, is recording the homepage information described in HTML etc., and reads and provides an applicable page according to the demand from the information processor 12.

[0034]Drawing 3 is a figure showing the detailed example of composition of the information processor 12 shown in drawing 2. the voice response part 13 which divides the information processor 12 roughly and performs processing between the telephones 10 as shown in this figure, and the browsing part 14 which performs control which downloads homepage information -- and, It is constituted by the HTML analyzing parts 15 which analyze the downloaded homepage information.

[0035]Here, the voice response part 13 is constituted by the voice recognition part 13a, the dial recognition part 13b, and the speech synthesis section 13c. The voice recognition part 13a recognizes the audio signal from the telephone 10, and notifies a recognition result to the telephone-operations analyzing parts 14a.

[0036]When the dial of the telephone 10 is operated, the dial recognition part 13b recognizes the contents of operation, and notifies them to the telephone-operations analyzing parts 14a. The speech synthesis section 13c changes into an applicable audio signal the text information supplied from the outputting part 15d according to control of the sound reproduction control section 14b, and transmits to the telephone 10 via the public network 11.

[0037]The browsing part 14 is constituted by the telephone-operations analyzing

parts 14a, the sound reproduction control section 14b, the hyperlink control section 14c, and the control section 14d in the same URL.

[0038]The telephone-operations analyzing parts 14a analyze the demand which the user performed by a sound or dial operation, and notify the analysis result to the sound reproduction control section 14b, the hyperlink control section 14c, and the control section 14d in the same URL.

[0039]The sound reproduction control section 14b controls the voice synthesis processing by the speech synthesis section 13c. The hyperlink control section 14c requires transmission of predetermined homepage information from the server 17.

[0040]In the same homepage, the control section 14d in the same URL controls movement in the following line from a predetermined line, movement in other paragraphs, etc. The HTML analyzing parts 15 are constituted by the component analyzing parts 15a, the barrier-free HTML analyzing parts 15b, the substitution part 15c, and the outputting part 15d.

[0041]The component analyzing parts 15a analyze the component of homepage information. The barrier-free HTML analyzing parts 15b analyze this, when the tag as HTML extended to homepage information is contained, and they notify an analysis result to the component analyzing parts 15a.

[0042]The substitution part 15c performs processing which replaces the predetermined information included in homepage information using other information. The outputting part 15d supplies the homepage information supplied from the component analyzing parts 15a and the substitution part 15c to the speech synthesis section 13c.

[0043]Next, operation of an embodiment of the invention is explained. After the information processor 12 closes a circuit in response to the call from the telephone 10 and drawing 4 performs predetermined processing, it is a flow chart with which the flow of processing is explained until it closes a circuit. A start of this flow chart will perform the following processings.

[S1] When the call from a user is received (receipt was carried out), the information processor 12 progresses to Step S2, and when other, it repeats the same processing.

[S2] The telephone-operations analyzing parts 14a receive the password transmitted, for example when a user operated a dial, and perform user authentication of whether to be a valid user.

[0044]This authenticating processing can also be omitted.

[S3] When the voice recognition part 13a inputs a user's sound, it progresses to step S4, and in being other, it repeats the same processing.

[S4] The voice recognition part 13a performs recognition processing of the sound which received.

[S5] The browsing part 14 performs processing applicable to the recognition result in step S4. When the connection request to a predetermined homepage is made, specifically, download of the homepage to which the hyperlink control section 14c corresponds is performed.

[S6] The information processor 12 judges whether processing is ended or not, when ending, it progresses to Step S7, and in being other, it returns to Step S3 and repeats the same processing.

[0045] For example, it progresses to Step S7 noting that processing is ended, when it judges whether the user did on hook [ of the hand set of the telephone 10 ], and cut the communication line, and it cuts.

[S7] The information processor 12 performs processing which cuts the communication line closed between the telephones 10.

[0046] According to the above processing, it becomes possible for a user to make the information processor 12 perform processing according to the contents which spoke via the telephone 10, and the contents of operation of the dial. Next, operation of a 1st embodiment of this invention is explained with reference to drawing 5. At a 1st embodiment of this invention, contained-text information is extracted by a block unit to homepage information, and each block is sound-ized in an order according to the attribution information of each block. A start of this flow chart will perform the following processings.

[S20] The hyperlink control section 14c requires the predetermined homepage information which the user demanded with the sound of the server 17.

[0047] For example, it is required that the hyperlink control section 14c should acquire homepage information as shown in drawing 6 from the server 17 according to the demand from a user. In this example, the table which consists of three cells is displayed on the viewing area 50a of the window 50, and the left end cell is further divided into the cell of two upper and lower sides. "An election public announcement and 1400 candidate persons" are displayed on the cell 50b of the left end upper part, and the cell 50c under it was notified publicly of "election. 4:00 p.m. ..." is displayed. A picture is displayed on a central cell and the menu item is listed by the right end cell 50d.

[0048][S21] The component analyzing parts 15a input the homepage information transmitted according to the demand of the hyperlink control section 14c, and develop it in the element as an element surrounded with the tag. Since an element shows the



lump of the meaning in an HTML document and the element has a layered structure, when a suitable hierarchy divides an HTML document, each element becomes equal to a block.

[0049] Drawing 7 is a figure showing an example of the HTML document corresponding to drawing 6. As shown in this figure, the tag and text information which show an attribute are put together, and the HTML document corresponding to the homepage shown in drawing 6 is formed.

[0050] The component analyzing parts 15a develop such homepage information for every element. In the example of drawing 6, "an election public announcement, candidacy ...", and "election were announced publicly here. Since 4:00 p.m. ..." and "North-South summit meeting ..." are stored in a cell different, respectively, it is recognized as a separate element.

[S22] The component analyzing parts 15a judge whether the processing to all the elements was completed, and when it completes, they end processing. In being other, it progresses to Step S23.

[S23] The component analyzing parts 15a analyze the tag contained in the HTML document.

[0051] That is, the component analyzing parts 15a analyze attribution information, such as a font size currently embedded to the tag belonging to each element. In the case of the title "an election public announcement and candidacy ..." shown in drawing 6, since the font size is 6 (font size="6") as shown in the 6th line of drawing 7, specifically, this is acquired.

[S24] The component analyzing parts 15a analyze a contained text in each element.

[0052] That is, the component analyzing parts 15a analyze a contained text in each element, and acquire the number of characters.

[S25] According to the analysis result of Step S23 and Step S24, the component analyzing parts 15a determine the read-aloud ranking of the element, and store it in an element table. Here, in the example of drawing 6, since the font size of a title "an election public announcement and candidate ..." is the maximum in "6", the ranking of this title is set to "1." A text "the election was announced publicly. Since the font sizes of 4:00 p.m. ..." and a menu "North-South summit meeting ..." are "3" and "2", respectively, each ranking is set to "2" and "3."

[0053] Drawing 8 is an example of the element table corresponding to the homepage shown in drawing 6. In this example, each element which constitutes the example of drawing 6 is stored with ranking. For example, the 1st item supports the title at the upper left of drawing 6 "an election public announcement and candidate ----", and the

ranking is set to "1."

[0054]A read-aloud order according to the attribute of the text information included in the element by the above processing to the element which constitutes homepage information is given. Thus, the generated element table is supplied to the speech synthesis section 13c via the outputting part 15d with text data.

[0055]The speech synthesis section 13c performs processing which sound-izes the supplied text data with reference to an element table. In the present example, "the election public announcement and candidacy" whose ranking is a title of the cell 50b of the left end shown in drawing 6 which is "1" are first sound-ized in an element table. [ ... ] Then, the cell 50c under it "the election was announced publicly. The present ..." will be sound-ized at 4:00 p.m., and, finally "- North-South summit meeting ..." of right end cell [ 50d ] will be sound-ized.

[0056]As explained above, according to the embodiment of the invention, the text information which constitutes a homepage, Since it divides per block (element) and the turn of sound-izing was determined as each block according to the attribute of contained-text information, it becomes possible to give priority to and sound-ize important information.

[0057]In an above embodiment, although a sound-ized order of each block was determined on the basis of the font size, it may be made to be based to a block on the number of characters and others of a contained text, for example.

[0058]Although contained-text information was treated as one block (element) in the cell of the table, it may be made to consider in an above embodiment that not only a method such but the text information arranged adjacently, for example is one block.

[0059]Next, a 2nd embodiment of this invention is described. Drawing 9 is a flow chart explaining operation of a 2nd embodiment of this invention. Since the composition of a 2nd embodiment is the same as that of the case of a 1st embodiment, the explanation is omitted.

[0060]A start of this flow chart will perform the following processings.

[S40] The hyperlink control section 14c requires the predetermined homepage information which the user demanded with the sound as acquiring from the server 17.

[0061]For example, it is required that the hyperlink control section 14c should acquire homepage information as shown in drawing 10 from the server 17 according to the demand from a user. In this example, while the title "hit chart of this week" is displayed, the track name which carried out rank yne to the hit chart, and the singer's name are displayed from primacy to the 4th place.

[S41] The component analyzing parts 15a input the homepage information transmitted

according to the demand of the hyperlink control section 14c, and develop it in the element as an element surrounded with the tag.

[0062] That is, the component analyzing parts 15a acquire the HTML document of the homepage used as an object as shown in drawing 11, extract the element (component surrounded with the tag) contained in this HTML document, and generate an element table.

[0063] Drawing 12 is a figure showing an example of the generated element table. In this example, the data contained in each element is stored with an element number, size, an attribute, etc.

[S42] The substitution part 15c reads a file substitution table applicable from the storage parts store which is not illustrated while receiving an analysis result from the component analyzing parts 15a.

[0064] Here, a file substitution table is a table for replacing the information included in the homepage by other information, or adding new information to the information included in the homepage, and the example is shown in drawing 13. In this example, the file name of the file on which the music corresponding to each element shown in drawing 12 was recorded matches, and is stored in the element number. For example, the replacing data corresponding to an element number "ELE001" is "grand\_son.mp3", and it turns out that the file format is "MP3" form from an extension.

[S43] The substitution part 15c replaces the data of an applicable element with reference to a file substitution table.

[0065] For example, in the case of the element [ 1st ], the data "1 place GRAND SON ..." contained in the element is replaced by grand\_son.mp3 which is music data of MP3 form. It is also possible to add the file of MP3 form, without replacing.

[S44] While the speech synthesis section 13c receives supply of the homepage information by which substitution was made and sound-izes text information, when there is music data newly added by substitution, it reproduces the music data.

[0066] Therefore, when the homepage shown in drawing 10 by which substitution was made by the above processing is sound-ized, First, after "the hit chart of this week" currently displayed on the viewing area 60a of the window 60 is uttered, the music which is carrying out rank yne to the 4th place from the primacy of the hit chart will be reproduced in order.

[0067] Although some data contained in an element was replaced by music data in an above embodiment, it may be made to add music data to the end of an element, without replacing. According to such composition, after "he is GRAND SON Akira Asada the 1st place" is uttered, the musical file "grand\_son.mp3" will be reproduced.



[0068]In an above embodiment, when a file substitution table was prepared per homepage and homepage information was acquired from a server, acquire an applicable file substitution table, and were made to perform substitution processing, but. It may be made to direct the data used as the candidate for substitution with the extension tag of HTML.

[0069]For example, a HTML tag is extended, a barrier-free tag <BF> is defined, and it may be made to embed the file name of the file which it is going to add to this tag. The barrier-free tag embedded by doing in this way is disregarded as a meaningless tag, when perusing by the usual browser, but. As opposed to the information processor 12 as a portal, the barrier-free HTML analyzing parts 15b are provided, and a barrier-free tag <BF> can be analyzed by these barrier-free HTML analyzing parts 15b.

[0070]Although text information was replaced by speech information in an above embodiment, For example, picture information is replaced by direct sound voice information, or picture information is replaced by applicable text information (for example, text information explaining the picture), and it may sound-be made toize by the speech synthesis section 13c.

[0071]Finally, the above-mentioned processing capability is realizable by computer. In that case, the contents of processing of the function which an information processor should have are described by computer by the program recorded on the recording medium which can be read, and the above-mentioned processing is realized by the computer by executing this program by computer. As a recording medium which can be read, there are a magnetic recording medium, semiconductor memory, etc. by computer. In making it circulate to a commercial scene, store and circulate a program to portability type recording media, such as CD-ROM (Compact Disk Read Only Memory) and a flexible disk, or, It stores in the storage device connected via the network, and can also transmit to other computers through a network. When performing by computer, the program is stored in the hard disk drive in a computer, etc., and it loads to main memory and performs.

[0072]

[Effect of the Invention]In the information processor which performs processing which reads the homepage information described to have explained above by the markup language by this invention, and with which a user is provided, A call receptionist means to receive the call by the telephone from a user, and the voice recognition means which recognizes the sound of the user transmitted from telephone, When it has been recognized that the demand which accesses a predetermined homepage by a voice recognition means was made, The homepage information acquisition means which

acquires corresponding homepage information, The text information extracting means which extracts contained-text information by a block unit to homepage information, The attribute information extracting means which extracts the attribution information of each block extracted by the text information extracting means, The sound-ized order determination means which determines an order of sound-izing of each block according to the attribution information of each block extracted by the attribute information extracting means, Since it was made to have a sound-ized means to sound-ize the text contained in each block, based on the decision results of a sound-ized order determination means, When sound-izing homepage information, it becomes possible to express the information on the homepage currently expressed by attributes, such as a size of a character, by an utterance order.

[0073]In the information processor which performs processing which reads the homepage information described by the markup language, and with which a user is provided, A call receptionist means to receive the call by the telephone from a user, and the voice recognition means which recognizes the sound of the user transmitted from telephone, When it has been recognized that the demand which accesses a predetermined homepage by a voice recognition means was made, The homepage information acquisition means which acquires corresponding homepage information, Since it was made to have a permutation means which replaces the predetermined information included in homepage information by other information, and a sound-ized means to sound-ize homepage information by which the substitution of information was made by the permutation means, For example, lack of information can be prevented when sound-izing homepage information by replacing picture information by applicable speech information.

---

[Translation done.]

---

## DESCRIPTION OF DRAWINGS

---

[Brief Description of the Drawings]

[Drawing 1]It is a principle figure explaining the principle of operation of this invention.

[Drawing 2]It is a figure showing the example of composition of an embodiment of the invention.

[Drawing 3] It is a figure showing the detailed example of composition of the information processor shown in drawing 2.

[Drawing 4] In the embodiment shown in drawing 3, it is a flow chart for explaining the flow of processing from the receipt from a user to cut treating.

[Drawing 5] It is a flow chart for explaining the flow of the processing for determining a sound-sized order of a block.

[Drawing 6] The flow chart shown in drawing 5 is an example of the homepage made into the object of processing.

[Drawing 7] It is an HTML document corresponding to the homepage shown in drawing 6.

[Drawing 8] It is an example of the element table obtained when the processing shown in drawing 5 is applied to the homepage shown in drawing 6.

[Drawing 9] It is a flow chart for explaining operation of a 2nd embodiment of this invention.

[Drawing 10] The flow chart shown in drawing 9 is an example of the homepage made into the object of processing.

[Drawing 11] It is an HTML document corresponding to the homepage shown in drawing 10.

[Drawing 12] It is a figure showing an example of an element table.

[Drawing 13] It is a figure showing an example of a file substitution table.

[Description of Notations]

1 Information processor

1a Call receptionist means

1b Voice recognition means

1c Homepage information acquisition means

1d text information extracting means

1e Attribute information extracting means

1 f Sound-sized order determination means

1 g Sound-sized means

2 Public network

3 Telephone

4 Internet

5 Server

10 Telephone

11 Public network

12 Information processor



- 13 Voice response part
  - 13a Voice recognition part
  - 13b Dial recognition part
  - 13c Speech synthesis section
- 14 Browsing part
  - 14a Telephone-operations analyzing parts
  - 14b Sound reproduction control section
  - 14c Hyperlink control section
  - 14 d Control section in the same URL
- 15 HTML analyzing parts
  - 15a Component analyzing parts
  - 15b Barrier-free HTML analyzing parts
  - 15c Substitution part
  - 15 d Outputting part
- 16 Internet
- 17 Server

---

[Translation done.]